Assignment 2: Pattern Recognition

IIT Jammu

February 7, 2025

Submission Guidelines

- The assignment is due on **February 20, 2025, at 11:59 PM**.
- Ensure that your code is well-commented and structured.
- Visualizations should be properly labeled and explained.
- Plagiarism will result in deduction of marks.

Question 1: Generating Gaussian Distributed Data

1. Consider three Gaussian distributions, namely $N(\mu_1, \Sigma_1)$, $N(\mu_2, \Sigma_2)$, and $N(\mu_3, \Sigma_3)$, where μ_1, μ_2, μ_3 are the mean vectors, and $\Sigma_1, \Sigma_2, \Sigma_3$ are the covariance matrices of the distributions.
2. Generate 50 random data points from $N(\mu_1, \Sigma_1)$, 40 data points from $N(\mu_2, \Sigma_2)$, and 60 data points from $N(\mu_3, \Sigma_3)$. Label them as D_1, D_2, D_3, belonging to class 1, class 2, and class 3, respectively.
3. Shuffle the dataset and remove the class labels.
4. Visualize the dataset using scatter plots.

Question 2: Implementing Clustering Algorithms

1. Using the dataset generated in Question 1, apply the following clustering algorithms to group the data into three clusters:
   (a) **Gaussian Mixture Model (GMM)** - Implement and visualize the clusters.
   (b) **K-Means Clustering** - Apply K-means with $k = 3$ and analyze the results.
   (c) **Agglomerative Hierarchical Clustering** - Perform hierarchical clustering and display the dendrogram.

Question 3: Discriminant Function and Multivariate Normal Distribution

1. Given the following dataset with feature vectors $X = (x_1, x_2)$ belonging to three classes $\omega_1, \omega_2, \omega_3$:

	ω_1	ω_1	ω_1	ω_1	ω_2	ω_2	ω_2	ω_2	ω_3	ω_3	ω_3	ω_3
x_1	2.1	1.1	1.4	3.3	4.4	3.4	4.5	4.1	-1.3	-3.2	-3.2	-2.1
x_2	-2.5	-3.1	-2.1	-1.8	6.5	5.8	7.2	5.65	-2.3	-4.5	-4.5	-3.3

2. Compute the mean vector μ over all samples, and class means μ_1, μ_2, μ_3.

3. Use `numpy.cov()` to compute the covariance matrices $\Sigma_1, \Sigma_2, \Sigma_3$ for each class.

4. Out of the three cases, determine which case applies for computing the discriminant function of the Multivariate Normal Density Function (NDF).

5. Let $P(\omega_1) = 0.4$, $P(\omega_2) = 0.35$, and $P(\omega_3) = 0.25$. Write a Python function to compute the discriminant functions defined in part (c).

6. Compute and plot the discriminant functions $g_1(X), g_2(X), g_3(X)$ along with the sample points in 2D space.